

1645

RECEIVED

APR 30 2002

TECH CENTER 1600/2900



ENTERED

1600

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/895,913A

DATE: 04/18/2002

TIME: 14:48:30

Input Set : A:\06132.043002.SEQLIST.TXT

Output Set: N:\CRF3\04182002\I895913A.raw

RECEIVED

MAY 15 2002

TECH CENTER 1600/2900

4 <110> APPLICANT: Kleanthous, Harold
 5 Al-Garawi, Amal
 6 Miller, Charles
 7 Tomb, Jean Francois
 8 Oomen, Raymond P.
 10 <120> TITLE OF INVENTION: Identification of Polynucleotides
 11 Encoding Novel Helicobacter Polypeptides in the Helicobacter
 12 Genome

14 <130> FILE REFERENCE: 06132/043002
 16 <140> CURRENT APPLICATION NUMBER: US 09/895,913A
 17 <141> CURRENT FILING DATE: 2001-06-29
 19 <150> PRIOR APPLICATION NUMBER: US 08/881,227
 20 <151> PRIOR FILING DATE: 1997-06-24
 22 <160> NUMBER OF SEQ ID NOS: 368
 24 <170> SOFTWARE: FastSEQ for Windows Version 4.0
 26 <210> SEQ ID NO: 1
 27 <211> LENGTH: 1004
 28 <212> TYPE: DNA
 29 <213> ORGANISM: Helicobacter pylori
 31 <220> FEATURE:
 32 <221> NAME/KEY: CDS
 33 <222> LOCATION: (56)...(931)
 35 <400> SEQUENCE: 1

36 cattaaacgc atgattttg ctacaataat aggatttaa ttatataaag gacaa atg 58
 37 Met
 38 1
 40 ggc atg cca aat agg ggc gtt gtt tta tta gac ggg caa gcg cta gct 106
 41 Gly Met Pro Asn Arg Gly Val Val Leu Leu Asp Gly Gln Ala Leu Ala
 42 5 10 15
 44 gat aat ata gaa aaa gat ttg aaa cat aaa atc caa ata ata acc gca 154
 45 Asp Asn Ile Glu Lys Asp Leu Lys His Lys Ile Gln Ile Ile Thr Ala
 46 20 25 30
 48 caa acg cat aaa cgc ccc aaa cta gcc gtg att tta gtg ggg aaa gat 202
 49 Gln Thr His Lys Arg Pro Lys Leu Ala Val Ile Leu Val Gly Lys Asp
 50 35 40 45
 52 ccc gct agt atc act tat gtc aat atg aag atc aaa gca tgc gaa agg 250
 53 Pro Ala Ser Ile Thr Tyr Val Asn Met Lys Ile Lys Ala Cys Glu Arg
 54 50 55 60 65
 56 gtg ggc atg gat ttt gac tta aaa acc ctc caa gaa aat att act gaa 298
 57 Val Gly Met Asp Phe Asp Leu Lys Thr Leu Gln Glu Asn Ile Thr Glu
 58 70 75 80
 60 gcc aaa ttg cta tcc ttg att aaa gat tac aat acc gat caa aac att 346
 61 Ala Lys Leu Leu Ser Leu Ile Lys Asp Tyr Asn Thr Asp Gln Asn Ile

RECEIVED
MAY 14 2002
TC 1700

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/895,913A

DATE: 04/18/2002
TIME: 14:48:30

Input Set : A:\06132.043002.SEQLIST.TXT
Output Set: N:\CRF3\04182002\I895913A.raw

62	85	90	95	
64	tca ggc gtt tta gtc cag ctc cct ttg ccc aga cac att gat act aaa			394
65	Ser Gly Val Leu Val Gln Leu Pro Leu Pro Arg His Ile Asp Thr Lys			
66	100	105	110	
68	atg att tta gaa gcc att gac cca aac aaa gat gtg gat ggt ttc cac			442
69	Met Ile Leu Glu Ala Ile Asp Pro Asn Lys Asp Val Asp Gly Phe His			
70	115	120	125	
72	ccc ctt aat atc ggt aag ctc tgc act caa aaa gaa tcg ttt ctg cca			490
73	Pro Leu Asn Ile Gly Lys Leu Cys Thr Gln Lys Glu Ser Phe Leu Pro			
74	130	135	140	145
76	gcc acc cct atg ggc gtg atg cgg ctt tta gag cat tac cat att gaa			538
77	Ala Thr Pro Met Gly Val Met Arg Leu Leu Glu His Tyr His Ile Glu			
78	150	155	160	
80	atc aag ggt aag gat gtg gcg att att gga gcg agc aat atc att ggc			586
81	Ile Lys Gly Lys Asp Val Ala Ile Ile Gly Ala Ser Asn Ile Ile Gly			
82	165	170	175	
84	aaa cct tta agc atg ctc atg cta aac gct ggg gct agc gtg agc gtg			634
85	Lys Pro Leu Ser Met Leu Met Leu Asn Ala Gly Ala Ser Val Ser Val			
86	180	185	190	
88	tgc cat att ttg act aaa gac att agt ttt tac acc caa aac gct gat			682
89	Cys His Ile Leu Thr Lys Asp Ile Ser Phe Tyr Thr Gln Asn Ala Asp			
90	195	200	205	
92	att gtc tgc gtg ggc gtg ggt aaa cct gat ttg att aaa gcg agc atg			730
93	Ile Val Cys Val Gly Val Gly Lys Pro Asp Leu Ile Lys Ala Ser Met			
94	210	215	220	225
96	tta aaa aaa ggg gct gta gtg gtg gat att ggg atc aat cat ttg aac			778
97	Leu Lys Lys Gly Ala Val Val Val Asp Ile Gly Ile Asn His Leu Asn			
98	230	235	240	
100	gat ggg cgt atc gtg ggc gat gtg gat ttt aac aac gtg caa aaa gtc			826
101	Asp Gly Arg Ile Val Gly Asp Val Asp Phe Asn Asn Val Gln Lys Val			
102	245	250	255	
104	gcc ggt ttt atc acc cct gtg cct aaa ggc gtg ggg cct atg acg att			874
105	Ala Gly Phe Ile Thr Pro Val Pro Lys Gly Val Gly Pro Met Thr Ile			
106	260	265	270	
108	gtc tcg ctt tta gaa aac act cta atc gct ttt gaa aaa caa caa agg			922
109	Val Ser Leu Leu Glu Asn Thr Leu Ile Ala Phe Glu Lys Gln Gln Arg			
110	275	280	285	
112	aag gga ttt taatgaaatt ttacgctct gtttatgcat ttgcctccag			971
113	Lys Gly Phe			
114	290			
116	ttgggttaggg acgattgtta ttgtgctgtt ggt			1004
118	<210> SEQ ID NO: 2			
119	<211> LENGTH: 292			
120	<212> TYPE: PRT			
121	<213> ORGANISM: Helicobacter pylori			
123	<400> SEQUENCE: 2			
124	Met Gly Met Pro Asn Arg Gly Val Val Leu Leu Asp Gly Gln Ala Leu			
125	1	5	10	15
126	Ala Asp Asn Ile Glu Lys Asp Leu Lys His Lys Ile Gln Ile Ile Thr			

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/895,913A

DATE: 04/18/2002
TIME: 14:48:30

Input Set : A:\06132.043002.SEQLIST.TXT
Output Set: N:\CRF3\04182002\I895913A.raw

127 20 25 30
128 Ala Gln Thr His Lys Arg Pro Lys Leu Ala Val Ile Leu Val Gly Lys
129 35 40 45
130 Asp Pro Ala Ser Ile Thr Tyr Val Asn Met Lys Ile Lys Ala Cys Glu
131 50 55 60
132 Arg Val Gly Met Asp Phe Asp Leu Lys Thr Leu Gln Glu Asn Ile Thr
133 65 70 75 80
134 Glu Ala Lys Leu Leu Ser Leu Ile Lys Asp Tyr Asn Thr Asp Gln Asn
135 85 90 95
136 Ile Ser Gly Val Leu Val Gln Leu Pro Leu Pro Arg His Ile Asp Thr
137 100 105 110
138 Lys Met Ile Leu Glu Ala Ile Asp Pro Asn Lys Asp Val Asp Gly Phe
139 115 120 125
140 His Pro Leu Asn Ile Gly Lys Leu Cys Thr Gln Lys Glu Ser Phe Leu
141 130 135 140
142 Pro Ala Thr Pro Met Gly Val Met Arg Leu Leu Glu His Tyr His Ile
143 145 150 155 160
144 Glu Ile Lys Gly Lys Asp Val Ala Ile Ile Gly Ala Ser Asn Ile Ile
145 165 170 175
146 Gly Lys Pro Leu Ser Met Leu Met Leu Asn Ala Gly Ala Ser Val Ser
147 180 185 190
148 Val Cys His Ile Leu Thr Lys Asp Ile Ser Phe Tyr Thr Gln Asn Ala
149 195 200 205
150 Asp Ile Val Cys Val Gly Val Gly Lys Pro Asp Leu Ile Lys Ala Ser
151 210 215 220
152 Met Leu Lys Lys Gly Ala Val Val Val Asp Ile Gly Ile Asn His Leu
153 225 230 235 240
154 Asn Asp Gly Arg Ile Val Gly Asp Val Asp Phe Asn Asn Val Gln Lys
155 245 250 255
156 Val Ala Gly Phe Ile Thr Pro Val Pro Lys Gly Val Gly Pro Met Thr
157 260 265 270
158 Ile Val Ser Leu Leu Glu Asn Thr Leu Ile Ala Phe Glu Lys Gln Gln
159 275 280 285
160 Arg Lys Gly Phe
161 290
164 <210> SEQ ID NO: 3
165 <211> LENGTH: 2162
166 <212> TYPE: DNA
167 <213> ORGANISM: Helicobacter pylori
169 <220> FEATURE:
170 <221> NAME/KEY: CDS
171 <222> LOCATION: (66)...(2099)
173 <400> SEQUENCE: 3
174 cccttgcttc ttttgtctt ttttaagact ttatctcttg ttaaaaaaaag gttgtattaa 60
175 cgott atg aaa tcc cta tct aat gcc ctt ttt tcg ctc ttt tta aaa ggt 110
176 Met Lys Ser Leu Ser Asn Ala Leu Phe Ser Leu Phe Leu Lys Gly
177 1 5 10 15
179 ttt tat ttc acc ttt ttt atg agc ttg ttg ttt gtg ttt aat cgt atc 158
180 Phe Tyr Phe Thr Phe Phe Met Ser Leu Leu Phe Val Phe Asn Arg Ile

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/895,913A

DATE: 04/18/2002
TIME: 14:48:30

Input Set : A:\06132.043002.SEQLIST.TXT
Output Set: N:\CRF3\04182002\I895913A.raw

181	20	25	30	
183	ggc ttt atc ctt tat act ggc tat tat aag cat gct tta aaa aac cct			206
184	Gly Phe Ile Leu Tyr Thr Gly Tyr Tyr Lys His Ala Leu Lys Asn Pro			
185	35	40	45	
187	gtt ttt gat gaa atc atc aaa acc cta ttc aat gga gcc aga tat gat			254
188	Val Phe Asp Glu Ile Ile Lys Thr Leu Phe Asn Gly Ala Arg Tyr Asp			
189	50	55	60	
191	aat cgt gtg gtc tca agc tta gcg att ctt ttt atc atc atc ggg tta			302
192	Asn Arg Val Val Ser Ser Leu Ala Ile Leu Phe Ile Ile Gly Leu			
193	65	70	75	
195	ttg ggg tta ttt atc cct aaa cac caa acc aaa atg ctt aat att gtg			350
196	Leu Gly Leu Phe Ile Pro Lys His Gln Thr Lys Met Leu Asn Ile Val			
197	80	85	90	95
199	gcg tat ttt tct atc gct att atc ctg ttt tta aac att gca aac att			398
200	Ala Tyr Phe Ser Ile Ala Ile Leu Phe Leu Asn Ile Ala Asn Ile			
201	100	105	110	
203	gtt tat tat ggt att tat ggg aat gtg ttt gat gaa aat tta ttg gaa			446
204	Val Tyr Tyr Gly Ile Tyr Gly Asn Val Phe Asp Glu Asn Leu Leu Glu			
205	115	120	125	
207	ttt ttg cat gaa gac acg ctc acg att tta aaa atg agc ggg gaa tac			494
208	Phe Leu His Glu Asp Thr Leu Thr Ile Leu Lys Met Ser Gly Glu Tyr			
209	130	135	140	
211	cct att ttt tct agt ttt tca ctc ttt tta atc ctt agc gtt tta acc			542
212	Pro Ile Phe Ser Ser Phe Ser Leu Phe Leu Ile Leu Ser Val Leu Thr			
213	145	150	155	
215	tct ttt atc tat ttc aaa ctc caa aac gac ctt ttt aaa ccc aaa aat			590
216	Ser Phe Ile Tyr Phe Lys Leu Gln Asn Asp Leu Phe Lys Pro Lys Asn			
217	160	165	170	175
219	gct tat caa gcc gcc cac acc aaa ccc ctt aaa act ttc att tta ttt			638
220	Ala Tyr Gln Ala Ala His Thr Lys Pro Leu Lys Thr Phe Ile Leu Phe			
221	180	185	190	
223	gcg ctt ttt tcc ctc aca caa atg ttt tac att aac gcg caa ttg agt			686
224	Ala Leu Phe Ser Leu Thr Gln Met Phe Tyr Ile Asn Ala Gln Leu Ser			
225	195	200	205	
227	ttt gtg ggc gcg tct tta gat ctc agc ata gag cca gcc aaa gat cct			734
228	Phe Val Gly Ala Ser Leu Asp Leu Ser Ile Glu Pro Ala Lys Asp Pro			
229	210	215	220	
231	ttt tta atg aaa att acc ccc gga gcg ttt cgc aac ctt tat ctt tta			782
232	Phe Leu Met Lys Ile Thr Pro Gly Ala Phe Arg Asn Leu Tyr Leu Leu			
233	225	230	235	
235	gca cgc aat tac aga caa agc cat aac ctt aaa ttc agc gat ttt gct			830
236	Ala Arg Asn Tyr Arg Gln Ser His Asn Leu Lys Phe Ser Asp Phe Ala			
237	240	245	250	255
239	aaa gaa acg cct tta gaa gtg gcg aaa aat tat ttc cat ctt aaa gag			878
240	Lys Glu Thr Pro Leu Glu Val Ala Lys Asn Tyr Phe His Leu Lys Glu			
241	260	265	270	
243	aac cct tca aac aac ctc tat gag ttg cta act cag aca agc cgc aac			926
244	Asn Pro Ser Asn Asn Leu Tyr Glu Leu Leu Thr Gln Thr Ser Arg Asn			
245	275	280	285	

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/895,913A

DATE: 04/18/2002
TIME: 14:48:30

Input Set : A:\06132.043002.SEQLIST.TXT
Output Set: N:\CRF3\04182002\I895913A.raw

247 aat tcc aat caa acc att caa cat gtt ttt tat atc gtt tca gag tct	974
248 Asn Ser Asn Gln Thr Ile Gln His Val Phe Tyr Ile Val Ser Glu Ser	
249 290 295 300	
251 ttg agt tca tgg cat ttt gat cca aaa ttt gac gct ata ggg cta acg	1022
252 Leu Ser Ser Trp His Phe Asp Pro Lys Phe Asp Ala Ile Gly Leu Thr	
253 305 310 315	
255 agc gct tta caa gac ttg gtt aaa aaa gag cat gcc cac atg ctt tct	1070
256 Ser Ala Leu Gln Asp Leu Val Lys Lys Glu His Ala His Met Leu Ser	
257 320 325 330 335	
259 gct ttt att gaa agc gcc cca cgg acc gtt aaa agc cta gat gtc caa	1118
260 Ala Phe Ile Glu Ser Ala Pro Arg Thr Val Lys Ser Leu Asp Val Gln	
261 340 345 350	
263 atc aca ggc tta ccc tat atc aat gat aat aac tta gtc aat tca ggg	1166
264 Ile Thr Gly Leu Pro Tyr Ile Asn Asp Asn Asn Leu Val Asn Ser Gly	
265 355 360 365	
267 gtg atc ctc cct agc ttt cct atg gcg att ggc aat atc aca aaa act	1214
268 Val Ile Leu Pro Ser Phe Pro Met Ala Ile Gly Asn Ile Thr Lys Thr	
269 370 375 380	
271 ctg ggt tat aaa aac aac ttt tat tat ggg ggt agc ggg att tgg aac	1262
272 Leu Gly Tyr Lys Asn Asn Phe Tyr Tyr Gly Ser Gly Ile Trp Asn	
273 385 390 395	
275 aaa ctc act agt ttc acc aaa aaa caa ggt ttt cac gcc ctt tat ttc	1310
276 Lys Leu Thr Ser Phe Thr Lys Lys Gln Gly Phe His Ala Leu Tyr Phe	
277 400 405 410 415	
279 aat aac cat ctc tta gaa ttt gcc caa aac aag ccc tac cct aaa ccc	1358
280 Asn Asn His Leu Leu Glu Phe Ala Gln Asn Lys Pro Tyr Pro Lys Pro	
281 420 425 430	
283 ata gag agc aac tgg gga gtg cat gat aat att tta ttt gac tat att	1406
284 Ile Glu Ser Asn Trp Gly Val His Asp Asn Ile Leu Phe Asp Tyr Ile	
285 435 440 445	
287 tta gaa aac acc aac ccc cat gaa aaa act ttc agc atg gtc atg act	1454
288 Leu Glu Asn Thr Asn Pro His Glu Lys Thr Phe Ser Met Val Met Thr	
289 450 455 460	
291 tta agc aac cat gcg atc aaa aac gtg aat ctc aaa gcc ttt ggc gtg	1502
292 Leu Ser Asn His Ala Ile Lys Asn Val Asn Leu Lys Ala Phe Gly Val	
293 465 470 475	
295 cct tta gaa aaa atc caa caa ttt gtg gaa aaa acc ccc aaa tca gaa	1550
296 Pro Leu Glu Lys Ile Gln Gln Phe Val Glu Lys Thr Pro Lys Ser Glu	
297 480 485 490 495	
299 aat tta ccg gac gct aat tct tta ggg cat att tac tgg tat gac aaa	1598
300 Asn Leu Pro Asp Ala Asn Ser Leu Gly His Ile Tyr Trp Tyr Asp Lys	
301 500 505 510	
303 gta atc gtc agt ttc atc aaa aaa gcc agc caa aaa ttc cct aac tcg	1646
304 Val Ile Val Ser Phe Ile Lys Lys Ala Ser Gln Lys Phe Pro Asn Ser	
305 515 520 525	
307 ctt ttt atc atc aca ggg gat cat ttt gac agg agc tat gaa tac gct	1694
308 Leu Phe Ile Ile Thr Gly Asp His Phe Asp Arg Ser Tyr Glu Tyr Ala	
309 530 535 540	
311 aaa aac gat ttg tat atc att aaa tcc gtg ccg ctt att tta tat gcc	1742

RAW SEQUENCE LISTING ERROR SUMMARY DATE: 04/18/2002
PATENT APPLICATION: US/09/895,913A TIME: 14:48:31

Input Set : A:\06132.043002.SEQLIST.TXT
Output Set: N:\CRE3\04182002\1895913A.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:13; Xaa Pos. 187
Seq#:14; Xaa Pos. 187
Seq#:33; N Pos. 976
Seq#:33; Xaa Pos. 309
Seq#:34; Xaa Pos. 309
Seq#:41; N Pos. 43
Seq#:59; N Pos. 331
Seq#:85; N Pos. 43
Seq#:146; Xaa Pos. 156
Seq#:307; N Pos. 71
Seq#:317; N Pos. 116
Seq#:317; Xaa Pos. 22
Seq#:318; Xaa Pos. 22
Seq#:361; N Pos. 738
Seq#:362; Xaa Pos. 224

VERIFICATION SUMMARY
PATENT APPLICATION: US/09/895,913A

DATE: 04/18/2002
TIME: 14:48:31

Input Set : A:\06132.043002.SEQLIST.TXT
Output Set: N:\CRF3\04182002\I895913A.raw

L:1886 M:258 W: Mandatory Feature missing, <223> not found for SEQ ID#:13
L:1886 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13 after pos.:690
L:1925 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:14 after pos.:176
L:3096 M:258 W: Mandatory Feature missing, <220> not found for SEQ ID#:33
L:3177 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:33 after pos.:968
L:3178 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:33 after pos.:1016
L:3268 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:34 after pos.:304
L:3565 M:258 W: Mandatory Feature missing, <220> not found for SEQ ID#:41
L:3566 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:41 after pos.:0
L:4638 M:258 W: Mandatory Feature missing, <220> not found for SEQ ID#:59
L:4660 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:59 after pos.:310
L:5937 M:258 W: Mandatory Feature missing, <220> not found for SEQ ID#:85
L:5938 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:85 after pos.:0
L:10191 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:146 after pos.:144
L:21688 M:258 W: Mandatory Feature missing, <220> not found for SEQ ID#:307
L:21690 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:307 after pos.:60
L:22087 M:258 W: Mandatory Feature missing, <220> not found for SEQ ID#:317
L:22096 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:317 after pos.:104
L:22097 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:317 after pos.:152
L:22183 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:318 after pos.:16
L:25506 M:258 W: Mandatory Feature missing, <220> not found for SEQ ID#:361
L:25564 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:361 after pos.:732
L:25607 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:362 after pos.:208